

Network Segmentation

By Tom Adams, Mark Wittry, Dorian Deane, Dwight Bynum, David Hayes, Hwa-Jung Han, Blair Miller, Dominick Garzaniti, and Tia Strombeck

Traditional network security has been based on separating the enterprise internal network from all external connections and controlling what is allowed to enter. This plan cannot deliver effective security in today's enterprise networks. It is based on the assumption that all the "good guys" are inside the company and all the "bad guys" are outside. In practice, this assumption is never true.

While bad guys sometimes manage to penetrate the castle defenses, most attacks are traced to the computers of insiders. Even with the best of intentions, any employee might be fooled by an e-mail, web page, or instant message containing a virus attachment—turning at least their computer, if not the employee, into a temporary bad guy.

If a traditional "perimeter defense" cannot protect today's enterprise, what can? Security-conscious enterprises are turning to fine-grained segmentation. This strategy, although focused on defense against cyber attacks, also pays dividends in terms of network flexibility and interactions with business partners. Segmented networks are orderly, carefully defined networks.

Defining Segmentation

Segmentation refers to dividing a network into distinct pieces so that communication between the segments may be controlled. This is not a new idea. For years, businesses have built networks with firewalls, effectively dividing the world into two network segments—inside the business and everywhere else. While that strategy is effective for the most basic of security technologies, the rising threat from viruses, trojans, and spyware necessitates a more granular approach.

Security segmentation is a control methodology, not a performance technology. Virtual Local Area Networks (VLANs), switches with multiple backplanes, and wireless networks with multiple ESSIDs¹ separate traffic to improve performance, not security. Segmentation for security purposes is achieved only if the network can prevent unauthorized communication between hosts. While these capabilities are sometimes found in VLANs, switches, and wireless access points, security segmentation features are most commonly embodied by devices like firewalls and routers with access control lists.

How Segmentation Protects

Segmentation limits communications to those exchanges that are required in the normal course of business. It protects against the abnormal, to the degree that the segmentation systems can distinguish between the two.

Viruses are a major threat to any enterprise. A virus can spread in minutes, far faster than any network security team can respond. Segmentation protects against this by limiting the scope of the spread. In a segmented network, an infected computer can communicate with only a few other systems, so the virus is slowed.

Another common threat is network intrusion. Intruders use many of the same techniques as virus writers. They have a small bag-of-tricks that lets them infiltrate a computer. They scan a large number of computers looking for systems where one of their "tricks" will work. Segmentation prevents communications outside the scope of normal business activity. Most of the intruders' scans are dropped by the network before they ever hit a potential target system.

Another benefit of segmentation is that company information leaks are limited even after a breach has occurred. Segmentation controls are applied equally to outbound and inbound network traffic. Even if an intruder makes their way into a vulnerable host, they have few options for shipping company secrets out.

Defense-in-Depth

All defenses can eventually be defeated. The only question is, “How much time and effort must the attacker devote to the job?” “Defense-in-depth” means using several layers of defensive measures to slow down and wear out an attacker, while at the same time alerting the defenders to the new threat. By applying several different protective systems at different points throughout the network, an attacker is forced to address each one individually. With diligence, the network security team can detect the attacker before he successfully breaches the final defense.

Islands of Protection

Segmentation divides an enterprise into multiple network “islands” and then strictly controls the communications between those islands. An island is defined as the area of the network that any given host can reach without having to pass through a network-based security checkpoint.² Often, an island is a subnet, protected by a security access control list (ACL) in a router.

Perfect protection is never achieved. There is always a balance between protection and efficiency. Ideally, we would like to have a router or firewall guarding each host, with a completely customized ACL rule set. Of course, this is not practical. We must group multiple hosts together in each island.

The hosts in an island can communicate with each other without passing through a security checkpoint. That means that we must be careful about how we group hosts. All the hosts in a given island should have similar communications requirements. That is, they should perform the same sorts of tasks and communicate in the same basic ways. A web server open to the public Internet should not be grouped into the same island as an internal web server, for example.

Hosts within an island also must be similar in their security requirements. Because these hosts can communicate directly, if one host in an island is compromised, it is likely that other hosts in that same island also will be compromised. Hosts with very different sorts of sensitive data should not be grouped together because that increases the risk that both types of information will be compromised.

Island Types

Verizon identifies five different island types based on their differing communications and security needs. There may be more than one island of each different type. An enterprise likely will have several public services islands, for example. This may be established based on security needs, to keep dissimilar servers from infecting each other in the event of a compromise. It also may be done for operational needs, such as a desire to have geographic diversity as a business continuity measure.

| Island Type | Purpose |
|---------------------------|--|
| Public Services | Landing and takeoff point for services available without regard to the IP address of the other endpoint. |
| Business-Partner Services | Systems that communicate to specific IP addresses or ranges outside the company. |
| Internal Applications | Services that only communicate directly with other systems inside the company. |
| Labs | These are test and development systems. |
| Desktops | These represent general office workers. |

Table 1: Island Types

Public Services

The public services islands encompass all servers that communicate to or from systems not controlled by the enterprise, without any consideration of the IP address of the external system. That is, the public servers offer their services to the entire Internet.³

Network firewalls do not protect public servers as fully as they guard servers in other sorts of islands. A firewall restricts packets based on the IP addresses and port numbers involved in the communications, but a firewall guarding a public services island cannot do this. Public servers are intentionally open to clients from any part of the Internet. Firewalls do offer some protection by restricting contact from the outside world only to those port numbers that are actually required for the service that is intended to be offered. A more sophisticated

firewall might also discard unexpected parameters and web page names outside of the scope of the application's normal needs.

Since public servers must accept traffic from the outside world, these applications likely will be attacked frequently. Some of these attacks will be successful. Both the applications and the public services islands should be designed to minimize damage in the event of a successful attack. The application's design should minimize the amount of sensitive data actually stored on a public services host.

Business-Partner Services

Business-partner service islands, like public services islands, offer services to hosts outside the control of the enterprise. The key difference is that these other hosts are known in advance. This allows the use of IP-layer filters, which drop all packets not destined to or from one of the pre-approved external hosts. The term "business partner" is used loosely to encompass any outside organization that the enterprise conducts business with, including vendors and suppliers, agents and marketing partners, and larger customers.

Data links to business partners may take many forms. Performance criteria may suggest private T1 or DS-3 links, or some form of virtual circuit service. Communications may be conducted over the open public Internet or via Private IP services. If the data that is to be exchanged is sensitive, encryption may be used to protect the privacy of the data while it crosses between the companies.

The design rules for business-partner service applications are somewhat relaxed. Compared to public-service applications, a business-partner server has greater non technical protections. The external systems are known in advance. Communication takes place within the structure of a defined business relationship. The external business has its own network and its own staff tasked with protecting that network. Contractual measures can be used to define the responsibilities each company has to the other in protecting their common systems and networks.

Internal-Application Islands

Internal-applications islands are the home of normal production hosts. These islands are where the day-to-day work of the company is done. Internal-applications islands never accept inbound service requests from any place external to the enterprise. They may be permitted to originate connections to external servers for defined purposes. For example, a server providing centralized patch management is an internal application, but it must contact software vendors outside the enterprise to download current patches and updates.

Labs

Labs need exceptional flexibility to support research and development. A lab might intentionally use old software to test how a system will behave for customers who don't have the latest software, or a virus could be deliberately introduced in a lab to test the attack resistance of an application.

Network segmentation around other island types exist to keep bad programs out and to slow them down if they do get in. Around lab islands, where bad software must be allowed to exist, the protective functions are designed to see that nothing escapes from the lab. Lab islands are generally prohibited from originating any outbound communication. Inbound communication is permitted from specific sources within the enterprise as required by the research.

Lab islands cannot be used to support production systems. Labs are isolated from the enterprise to the greatest degree possible, while production systems, by definition, must interact with the rest of the company.

Desktops

Desktop islands provide a home for the ordinary office worker's desktop (or laptop) PC. Desktop segments are not the proper home for most servers, but some general office devices that are technically considered servers are useful in these islands. Examples include network printers and Voice over IP (VoIP) handsets. An enterprise also may place local file servers in this area for performance reasons, but this should be carefully weighed against the additional risk posed by this choice.

Special Purpose Areas

Remote Access

Remote access is a generic term that includes all remote employee access. Some organizations use VPN systems based on IP Security. Others choose systems based on Secure Shell (SSH) or Secure Sockets Layer (SSL) technologies. Shops with an all Microsoft® environment may choose Point-to-Point Tunneling Protocol

(PPTP). Regardless of the underlying technology, remote access islands serve the same basic need—assisting an employee with authorized access to the company who needs to get work done from a remote location such as a hotel room or their home.

Remote access systems are a huge source of potential security vulnerabilities. They grant internal access to people outside the boundaries of the enterprise. These systems are subject to some of the most careful security measures.

Remote access is not a type of island. A remote access server logically straddles different sorts of islands to create a bridge between them.

There are three logical entry points to a remote access server. These may be implemented as physical interfaces or as logical entities within a single interface, but the essential point is that they serve different functions. Each of these entry points has characteristics of a different sort of island and receives different protections from the network infrastructure.

One entry point connects the remote server to the public Internet. Users connect to this entry point to request access. On this side, the remote-access server is a public services island with the most stringent restrictions possible. The remote access server accepts only those requests that are absolutely necessary to its function, but it must accept inbound requests from any global Internet address.

The second entry point connects the remote access server to the enterprise's internal network. On this side, the remote access server is another "desktop" island—just a collection of generic employee stations. This entry point gets the same restrictions as any other desktop island.

The third entry point is for actual management and control of the remote access server itself. In the first two entry points, packets are received to be passed through the remote access device and forwarded to somewhere else. On the management entry point, the packets actually are destined to the server itself.

The management entry point must be very carefully protected. Remote access servers intentionally create holes in the enterprise's network protections. An attacker with the ability to control a remote access server can misuse this power with devastating results. To prevent this, the remote access device management ports can only be reached from designated management systems. The remote access server also is permitted to contact the corporate user-authentication servers, time servers, and similar maintenance systems.

Wireless

Wireless access, based on IEEE 802.11 standards, is another remote access technology. It permits people who are outside the control of the corporate firewalls to access the corporate network. A wireless signal penetrates building walls, extending to nearby parking lots and public spaces. Wireless technologies must be regulated carefully so they will not become open portals to the "crackers" and "black hats" of the cyber world.

The wireless user must be considered a guest knocking at the door of the enterprise VPN server. They are unknown and are not trusted until they prove otherwise. Wireless users must be authenticated before they are allowed to access the corporate networks.

Wireless connections also must be encrypted. Snoopers can eavesdrop on information that was transmitted between the corporation and legitimate wireless users, even if the snoopers cannot send any messages of their own. If not protected by encryption, this information is easy prey for corporate spies.

These two requirements, authentication and encryption, were not sufficiently addressed by early generations of wireless standards. WEP, the Wired Equivalent Privacy protocol, can be broken by an attacker with a typical laptop PC and a few hours of spare time. Other newer protocols, such as WPA2 (IEEE-802.11i), will offer much better security in the future.

Despite the promise of improvement, today's wireless security is not adequate for enterprise use. As of March 2005, these newer protocols are not widely deployed and available yet, nor have they stood the test of time in actual use. The older WEP protocol has already failed that test. For this reason, inherent security features of wireless networks should not be relied upon to protect access to the enterprise.

Wireless can meet a legitimate need if deployed in combination with other security precautions. Wireless networks should be deployed as semi-public "hot spots."⁴ These provide a convenient means of connection for employees and visitors when in conference rooms and other areas where stringing wire is not a desirable practice. These wireless access points must be connected to an external Internet connection, not to the enterprise internal net. Using that Internet connection, employees can enter the company through the standard remote access VPN servers just as they would from a hotel room. Visitors can easily reach out to their own companies and their own remote access VPN systems.

Types of Security Controls

Network security controls can operate at several different layers in the communications hierarchy. Protections operating at higher protocol layers can be more discriminating because they examine more information. However, the most detailed protection mechanisms also require the greatest investment in management and computer resources to implement. For this reason we cannot solely rely on finely-tuned, host-based security on all systems. The workload to manage such a system would be overwhelming.

| ISO Layer | Protections | | |
|-----------------|---------------------------------------|-------------------|-----------------|
| 7. Application | Host-Based or Application Proxy | Stateful Firewall | Static Firewall |
| 6. Presentation | | | |
| 5. Session | | | |
| 4. Transport | | | |
| 3. Network | Future 802.11i | | |
| 2. Link | Physical Security (Except Wireless) | | |
| 1. Physical | Physical Security (Except Wireless) | | |

Table 2: ISO Layers and Associated Protections

Application-layer protections operate on the endpoint hosts at the top of the protocol stack. This type of protection has access to the broadest range of information, enabling careful discrimination between requests. At the applications layer, the very concept of a “packet” often has been discarded already. Application-layer protections operate by considering whole client requests and server responses, whether they are contained in a single network packet or span many packets.

Firewalls are less complex protections, operating mostly on the network and transport layers. These protections take no notice of the actual content being passed over a session. If a firewall permits traffic to port 80 (web requests) to pass through, it will permit communications if an attacker suddenly begins running Internet Relay Chat (IRC) sessions on port 80.⁵

Static firewalls evaluate each packet individually, using only the source and destination IP addresses and port numbers. They also sometimes treat the first packet of a TCP session differently, allowing for a limited form of session control. By contrast, stateful firewalls keep track of which network sessions are currently in use. A stateful firewall can base the decision to block a particular packet on not only the packet’s control headers, but also on the relationship of the current packet to other packets that have already passed.

This is most clearly illustrated in the case of UDP-based services, such as the Domain Name System (DNS). A static firewall knows only that DNS answers come from port 53. All traffic claiming a source port of 53 will be permitted to pass the firewall. A stateful firewall, on the other hand, remembers that a host has recently sent a DNS query to a particular server. It permits a reply packet from that server, coming from port 53, going back to the originating host, and only for a brief time. All other packets from UDP port 53 are disallowed.

Host Security Measures

While this paper focuses on segmentation’s place in network security, a wall is not built of a single brick. Regardless of the security features offered by the network infrastructure, the host computers must take an active part in their own defense.

Host defenses can implement applications-layer filters, basing decisions on information that is not available to a network-level defense. For example, a host-based application may filter requests based on the name of the program controlling the network session,⁶ or on the basis of the user login ID.⁷ User and program names are not part of the network or transport layers used by network-based defenses, so only host-based countermeasures have a chance to consider these factors.

A host should limit inbound packets to those coming from IP addresses and ports—if known—with which the server host has a business relationship. For some server types, this may mean permitting a very broad range of communications. An internal web server hosting an employee directory must accept packets from any company-internal IP address destined to the server’s own port 80. A departmental database server, on the other hand, may restrict itself to exchanging packets only with hosts on a specific desktop island. A server in the finance department could be very restrictive, accepting packets only from one or two mainframe billing hosts.

Filtering incoming packets is only half of the story. A host also should filter its outbound packets. As with incoming packets, a host should set its filters to the most restrictive configuration consistent with business needs.

Outbound filtering differs from inbound filters in that inbound filters only protect the particular host. Outbound filters are “good citizenship” measures protecting the rest of the network if the host is compromised. Early viruses were indiscriminate, blasting packets at all possible targets. Outbound filtering makes these follow-up target scans fail quietly without impacting network performance.

Outbound filters also protect the host itself. Today’s more sophisticated attack programs often create unseen communications channels back to their human “masters” so that the attack can be directed more precisely. Such attacks fail when they cannot “phone home.” This limits the damage to the attacked host, and prevents the progress of the attack to other systems entirely.

Network Security Measures

Management of the network security rules is a constant challenge. Network-based security measures act at the island boundaries, controlling the communications between different islands. Data exchanged between hosts in the same island is not subject to network-based controls. Network security provides the greatest control when using a large number of very small security islands, but a multitude of small islands greatly complicates the management task. Organizations must balance the need for greater control with the capabilities of their staff and their available resources.

The first duty of a network security device is to protect its island against unfriendly incoming packets. Each island has characteristic traffic patterns based on the island type and the particular hosts and services located there. An internal-services island, for example, might receive inbound web requests from other portions of the corporation. The island may not accept inbound Telnet at all, and it may permit SSH connections only from an authorized management area.⁸ No other inbound traffic is accepted except the port(s) used for the actual service provided.

The second duty of a network security device is to protect neighboring islands from malicious traffic originating locally. If a host on the island is compromised, an attacker may use that host as a stepping stone to exploit trust relationships with hosts on other islands. Random attacks also may be directed towards customers, business partners, or the Internet at-large. Network-based outbound controls can prevent this. Regardless of how completely an attacker controls an infected host, he cannot remove network-based controls. By filtering outgoing packets, network-based protection assures that further attacks never will leave the island.

Related Best Practices

Access List Definition

Segmentation is enabled by routers and subnets, but it is defined by the access control lists⁹ encountered between different segments. These access lists define which systems may speak to one another and regulate what they are permitted to say. ACLs are the heart of network-based defenses. Many host-based defenses also will be based on some form of ACL. The security of a segmented island network will be no stronger than the ACLs.

ACLs may be specified as a description of what is permitted (a white list) or as a list of what is excluded (a black list). Traditional intrusion detection systems operate with a black-list approach, sending out an alarm when known or suspected illicit traffic is present. This approach only detects known threats. For optimum security, islands should be separated with a white-list approach. Everything not on the list should be rejected.

ACLs should be defined in advance, before the hosts are actually placed into the network islands. It is tempting to avoid disruption during a migration by placing the hosts first and then defining the ACLs to accommodate whatever network activity the hosts exhibit. This approach does not create a secure network. All it does is recreate, in the new ACLs, the same vulnerabilities that exist already in the old hosts and applications. While the ACLs may be set loosely at first, in the end, a secure network must force its servers to follow secure communications rules.

Access List Management

Enterprises routinely devote great effort to defining the content of access lists. Equally important, though often overlooked, are the management processes and policies that support and define the ACLs.

Access list management begins by establishing clear responsibility. ACL control should be vested in a definite group with exclusive authority to define the firewall rules. Ideally, this security controls group is not the network operations group. This segregation reflects the differing goals of the two teams. The operations group is paid to make certain that traffic flows smoothly. The security controls group is charged with preventing the flow of malicious traffic.

The organization's network needs will change over time. New systems will come online and old servers will be retired. As these changes occur, the ACLs must be updated. The management process must capture the "who, what, when, and why" of each change. Rule changes must be documented, reviewed, and authorized. Most rules also should have an expiration date to assure that a rule does not outlive the business justification that gave rise to it.

Updated control assures an orderly change process. New access rules must be pushed out to the network in an orderly fashion, taking account of any ordering dependencies. Updated control also requires that the enterprise has a way to roll back a failed access list update.

Finally, ACL management includes configuration verification. Each security-enforcing device must be checked to ensure that its actual configuration matches its intended configuration. This is an ongoing process. A valid configuration check only means that the device has not been compromised yet. Depending on the importance of the routers and access lists, this validation may be performed more or less frequently.

Remote Systems Management

In any sizable enterprise, most server systems are managed remotely. The computers overwhelmingly outnumber the staff, and servers are installed in data centers away from the ordinary office environment. While we take great precautions to secure the servers, the staff manages them from normal PC systems without any extraordinary protections. Any compromise of an administrator's desktop PC system could impact the servers they manage.

To improve security, the server islands¹⁰ permit remote administrative logins only from a small set of centralized jump hosts or applications-layer proxy servers. To manage a remote server, an administrator first logs in to a jump host. The jump host carefully verifies the administrator's identity and authority then relays the administrator's session to the server to be managed.

This jump-host configuration presents many security advantages. By concentrating the jump hosts in a small handful of internal server islands, the jump hosts vastly simplify the firewall ACLs for all the other server islands; they no longer need to know where each administrator sits. A jump host can use the latest in two-factor authentication, even when the server ultimately being managed is a decade old and does not support any special authentication systems or protocols. Finally, since all remote administration is forced to pass through the jump hosts, they provide a natural place for any desired logging.

Security Testing

A security plan, like a fire emergency plan, is no good unless it is frequently tested. An enterprise should conduct its own security test before crackers, virus writers, or others do it for them. This is usually accomplished by a process of network scanning followed by automated and human-guided vulnerability testing.

A security tester encounters two problems in a segmented network that were not present in older network designs. The first problem is making the hosts reachable in order to scan them. Historically, with only external firewall security, scanning could be carried out from anywhere within an enterprise. The island approach makes that impossible because the network-based security devices prevent most packets from crossing between different islands. Host-based security precautions also may prevent scanning hosts that are in the same island as the scanning system. To permit scanning, some holes must be deliberately created in the security rules.

The second problem is testing the security access rules themselves. Creating holes in the security access lists may allow the scanners to test the hosts, but it does not tell us whether the segmentation rules are actually functioning. It's important to know how the network looks to an attacker who does not have special privileges.

These challenges may be addressed by creating two separate groups of scanning hosts, each in their own island. The first scanning group has no special privileges. For purposes of access rules in other islands, this scanning group is declared to be a desktop island. Most attacks come from a company's own employees being tricked into opening an e-mail attachment or clicking on a web link that they should have avoided. This triggers the attacker's software, which then runs on the employee's PC. Since that PC is in a desktop island, the scanning hosts also must be treated as desktops. This gives the scanners the same view of the enterprise's internal network that an attacker would have.

The second group of scanners also is an internal services island. However, this island is very special. As a matter of corporate policy, this island is given a mandatory "free pass" through all network- and host-based security access rules. Although it will appear to each individual system administrator that their security is being weakened by allowing unfettered access from this special scanning group, overall corporate security is improved. These scans will reveal out-of-date software, missing patches, and weak or default configurations. None of this information would be available without the free pass through the firewall access lists.

Specially-privileged machines represent a tempting target for any attacker, and the scanning group is no different. If a host in the privileged scanning group were to be compromised, it would be a powerful base from which to launch additional attacks throughout the enterprise. These scanners must be exceptionally well-protected.

The scanners must not accept any remote login except from the corporate management jump-host island. The jump hosts will restrict access, permitting the security group personnel to connect to the scanners only. Their identity should be verified with both a password and a hardware-based token.

Mere user authentication is not enough protection for such a powerful target. Inbound TCP service should be denied for all new sessions, except for an encrypted management or console session from the corporate jump hosts. Outbound scanner-initiated TCP sessions and their associated returning responses are acceptable. UDP traffic, which is harder to organize into sessions and police, should be blocked, except for DNS and NTP service from known corporate servers. All other network traffic should be blocked.

Migration

No enterprise can step immediately into the sort of network outlined herein. An enterprise is trapped by its own legacy networks. These networks have carried the load for years and sustain critical business operations. It is cost-prohibitive to simply abandon them and start over. The new island architecture must be introduced as a gentle evolution rather than an overnight shift.

Establish a Strong DNS System

Realigning servers into security islands will force many servers to change their IP addresses. A comprehensive DNS service helps prevent disruption during these changes. DNS names remain constant even while the servers are reassigned to new IP addresses and subnets. This removes the need for wholesale reconfiguration of other related systems each time a single host is moved.

Understand the Applications

The island concept is based on grouping hosts according to their security needs. But hosts don't have security needs, their applications do. Grouping the hosts into islands means understanding the applications that each host supports and the communications needs of each application.

As the islands are established, and the hosts moved into them, the network security systems can help identify traffic patterns. Applications often have many network dependencies that go unnoticed. Since the island security rules eventually will have a default, deny-everything-else ending, it is important to understand all the communications that an application may need.

By setting the traffic filter to log, rather than drop the unknown traffic, the undocumented exchanges can be discovered.

End Notes

1. The Extended Service Set ID (ESSID), sometimes called just the Service Set ID (SSID), is a name given to a wireless network. It serves to distinguish a specific wireless network from others operating in the same general area where their radio signals might overlap.
2. While we endorse and use host-based protections, we cannot rely on them. An intruder with physical access to a host can disable the host-based protections or even replace the entire host operating system.
3. This does not mean that everyone on the Internet can use any particular service. While the web server will deliver pages to anyone on the Internet, the web application may impose a login requirement, for example, or other access restrictions.
4. An access point is semi-public because, while not intended to provide service to the whole world, it must be treated for security purposes as though the whole world does, in fact, have access. In other words, we always treat an access point as though its radio network has already been compromised.
5. IRC is Internet Relay Chat. IRC is frequently used as a back door to provide remote control over armies of virus-infected machines, coordinating the sending of spam e-mail or launching of a distributed denial of service (DDOS) attack. IRC normally runs on TCP port 6667.
6. The host firewall, Zone Alarm®, offers this feature on Windows systems.
7. The Linux® IP tables host firewall modules can do this.
8. The management area itself is an internal services island, with its own traffic pattern.
9. Access control lists include not just traditional router ACLs, but also firewall rule sets and the configuration of other network-based controls.
10. These include public services, business-partner services, and internal-applications islands.

